



Smith, H., & Fotheringham, K. (2020). Artificial intelligence in clinical decision-making: Rethinking liability. *Medical Law International*, 1. <https://doi.org/10.1177/0968533220945766>

Peer reviewed version

Link to published version (if available):
[10.1177/0968533220945766](https://doi.org/10.1177/0968533220945766)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Sage Publications at <https://journals.sagepub.com/doi/10.1177/0968533220945766>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

ARTIFICIAL INTELLIGENCE IN CLINICAL DECISION-MAKING: RETHINKING LIABILITY

ABSTRACT

This article theorises, within the context of the law of England and Wales, the potential outcomes in negligence claims against clinicians and software development companies (SDCs) by patients injured due to AI system (AIS) use with human clinical supervision. Currently, a clinician will likely shoulder liability via a negligence claim for allowing defects in an AIS's outputs to reach patients. We question if this is 'fair, just and reasonable' to clinical users: we argue that a duty of care to patients ought to be recognised on the part of SDCs as well as clinicians. As an alternative to negligence claims, we propose 'risk pooling' which utilises insurance. Here, a fairer construct of shared responsibility for AIS use could be created between the clinician and the SDC; thus, allowing a rapid mechanism of compensation to injured patients via insurance.

KEYWORDS: artificial intelligence, duty of care, negligence, risk pooling, tort

I. ARTIFICIAL INTELLIGENCE AND HEALTHCARE

Decision-making for patients in the clinical environment has historically been led by the clinical professions. The development and limited introduction of AI systems (AISs) into the healthcare sector constitutes a novel, non-human aspect to clinical decision-making.¹ In this paper we consider AISs which are advisory rather than autonomous; they do not directly interact with patients, but are designed to aid and influence the clinician's thought processes. In the UK, the National Health Service (NHS) has started some preliminary collaborations with

¹ JL De Fauw *et al.*, 'Clinically applicable deep learning for diagnosis and referral in retinal disease' (2018) *Nature Medicine* 1342

software development companies (SDCs).² These collaborations indicate the possibility for AISs to be formally adopted into the UK's healthcare system at some point in the future. An AIS may be designed to learn from its experiences and adjust its outputs without being specifically programmed to do so ('machine learning').³ The process by which the system calculates its outputs could be sufficiently complex to effectively render it inscrutable to a non-expert user, a black box⁴ in common parlance. These characteristics increase the risk that an AIS could produce a clinically inappropriate recommendation and that the defective logic involved goes undetected. This risk necessitates some clinical oversight of the system's outputs to ensure the AIS's recommendations are safe and relevant to the patient.

There is currently a lack of clarity surrounding the sufficiency of legal mechanisms for liability when applied to malfunctioning AISs. In 2017a House of Lords Select Committee recommended that the Law Commission investigate whether current legal principles were adequate to address liability issues when using AI and to make recommendations in this area,⁵ but a formal reference has not yet been made by the Government. In lieu of these recommendations we perform legal analysis of these issues within the jurisdiction of England and Wales.

This paper will examine how the tort of negligence can be applied in the scenario where a clinician uses an AIS's inappropriate recommendation and the patient comes to harm as a result. First, we identify the conditions which put the clinician at risk of carrying the burden of claims in this scenario and how the SDC could be able to limit their liability. We will argue that this situation is unfair to the clinical user as the clinical decision-making space has been modified

² D Campbell, 'NHS to sign up patients for "virtual" A&E in tech revolution' The Guardian (London, 23 May 2019) <www.theguardian.com/society/2019/may/23/birmingham-to-begin-accident-and-emergency-online-chat-service-in-tech-revolution-for-nhs-care/> accessed 8 March 2020

³ J Turner, *Robot rules: regulating artificial intelligence* (Palgrave Macmillan 2019) 16

⁴ Mukherjee, S. 'A.I. Versus M.D.' *The New Yorker* 27 March 2017
<<https://www.newyorker.com/magazine/2017/04/03/ai-versus-md>>

⁵ Select Committee on Artificial Intelligence, *AI in the UK: ready, willing and able?* (HL 2017-19, 100) 135

by the SDC via their AIS. In the final sections of this paper we explore risk pooling as a possible solution which provides shared responsibility between SDCs and clinicians.

A. Key stakeholders

Clinicians may work as sole agents (e.g. a self-employed General Practitioner) as well as in groups (e.g. hospitals), often interchangeably. We use the term ‘clinician’ throughout to describe medically-trained staff working in either context. On the other hand, it would be extremely rare (but not impossible) for an AIS to be designed, developed, trained, tested, monitored, and updated by a single individual; software development is frequently undertaken by commercial enterprises. It is increasingly common to find interdisciplinary groups of technologists and clinicians working together. Here, we subsume clinicians involved in an interdisciplinary group as part of the SDC. However, at the point of care, clinicians are treated as a distinctive, independent professional grouping.

II. HARM

Whilst it might seem hard to envisage clinical professionals who lack specialist training heeding AISs which advise in specialist areas, this eventuality has happened, and its occurrence serves this paper as a useful vignette for discussion. IBM’s ‘Watson for Oncology’ ‘assesses information from a patient’s medical record and displays potential treatment options ranked by level of confidence, based on training with supporting evidence’.⁶ Watson was deployed at UB Songdo Hospital in Mongolia, where it was reported that its suggestions were followed at a rate of almost 100 per cent by clinicians who had little or no training in cancer care.⁷ The system was reported to have inappropriately recommended the drug Taxane for a patient whose

⁶ IBM (undated) *IBM Watson for Oncology: What Watson for Oncology can do for your organization* <<https://www.ibm.com/products/clinical-decision-support-oncology>> accessed 8 March 2020

⁷ C Ross and I Swetlitz, (2017) ‘IBM pitched Watson as a revolution in cancer care. It’s nowhere close.’ *STAT News*. <<https://www.statnews.com/2017/09/05/watson-ibm-cancer/>> accessed 8 March 2020

history would contraindicate the use of that drug.⁸ This error was noted by an oncology specialist,⁹ but it remains conceivable that a non-specialist clinician might not detect the error.¹⁰

III. CLINICAL NEGLIGENCE

Clinicians, and the hospitals in which they work, owe a duty of care to their patients.¹¹ Generally, patients are doubly vulnerable, by virtue of their health condition and relative lack of clinical knowledge. It remains the case that where a clinician has interpreted medical information and proceeds with a treatment plan that they have developed based on their clinical opinion, the clinician owes a duty of care towards their patient in the usual way.

The example of IBM's 'Watson' in Section II indicates that the clinician might rely upon an AIS to provide clinical recommendations. This poses a novel evidential problem for the claimant, as it is difficult to discern the relative influence of the human clinician and the non-human AIS. The clinician could instead be seen as a third party, a conduit for medical decisions that have been generated outside of the relationship of care that the clinician has with their patient.

Nevertheless, justice is not served by excluding claims where the clinician has chosen to use an AIS in providing treatment. To this end, some SDCs are adopting a posture that presents clinicians as the sole guardians of system safety. A description of IBM's Watson for Oncology system asserts that 'Watson does not make decisions on what a doctor should do. It makes recommendations based on hypothesis and evidence...'.¹² Thus, clinicians find themselves

⁸ *Ibid*

⁹ *Ibid*

¹⁰ The authors recognise that this scenario happened outside of England and Wales, but are using this as a historical example of AI system application in clinical decision-making which might be repeated or adopted in other regions. The authors are interested in the application of this scenario to the law of England and Wales as it is their home nation.

¹¹ *Barnet v Chelsea and Kensington Hospital Management Committee* [1969] 1 QB 428; *Darnley v Croydon Health Services NHS Trust* [2019] AC 831

¹² M Hengstler, E Enkel and S Duelli, 'Applied artificial intelligence and trust – the case of autonomous vehicles and medical assistance devices' (2016) 105 *Technological Forecasting and Social Change* 105, 115

trapped in a ‘moral crumple zone’¹³ where they become answerable for the AIS because of their choice to use the technology. Absorbing the moral responsibility for errors is a forceful argument for assigning full legal liability to clinicians, yet this is grossly unfair on clinicians as it disconnects accountability from the locus of control. Yet the clinician may act as an independent, knowledgeable intermediary between the software’s recommendations and the patient, but in practice is encumbered with the responsibility for computer-generated clinical advice over which they have only limited influence. It could be fairer for both the claimant and the clinician if the court entertained a form of joint liability, pooling the risk between the clinician and the SDC.

Two key cases underpin the consideration of clinical conduct in almost all negligent treatment and diagnosis claims: *Bolam v Friern Hospital Management Committee*¹⁴ and *Bolitho v City and Hackney Health Authority*.¹⁵ Clinical conduct is not usually considered negligent per *Bolam* if it is in accordance with a responsible body of opinion, and thus satisfies the standards of other responsible medical professionals. *Bolitho* requires that the standard relied upon has a logical basis. IBM promotes Watson for Oncology as a combination of the expertise of ‘leading oncologists’ in cancer care with ‘the speed of IBM Watson to help clinicians as they consider individualised cancer treatments for their patients’.¹⁶ The court has held in *Wilsher v Essex Area Health Authority*¹⁷ that the duty of care can be discharged by referring to a senior knowledgeable colleague for assistance. A clinician might not be successful in declaring that digitised expertise in the form of an AIS is equivalent to human proficiency, but they might be successful in arguing that a SDC had presented their AIS as dispensing reliable expert advice.

¹³ MC Elish, ‘Moral crumple zones: cautionary tales in human-robot interaction’ (2019) 5 *Engaging Science, Technology and Society* 40 <<https://estsjournal.org/index.php/ests/article/view/260>> accessed 8 March 2020

¹⁴ [1957] 2 All ER 118

¹⁵ [1998] AC 232

¹⁶ IBM (undated) IBM Watson for Oncology: What Watson for Oncology can do for your organization <<https://www.ibm.com/products/clinical-decision-support-oncology>> accessed 8 March 2020

¹⁷ [1987] QB 730 (CA); the case went to appeal in the House of Lords ([1988] AC 1074), but was concerned instead with causation and was silent on the question of referring to senior colleagues.

Thus, if an AIS such as Watson for Oncology is being presented by its SDC as dispensing expert advice, it would seem sensible for a clinician defendant to argue that in relying on the AIS, they are in fact acting in accordance with ‘a body of responsible medical opinion’.¹⁸

The law does not presume to differentiate between contradictory clinical opinions.¹⁹ For instance, an AIS may create a recommendation that might receive less than majority support among the medical community. Such a discovery could indeed represent a scientific advancement. Even if this was contestable, reliance on an unconventional recommendation for treatment would not necessarily be *a priori* negligent.²⁰ If the system’s outputs consisted of recommendations that were illogical, however, the claimant may be able to show that the duty of care had been breached by the clinician who acted on that recommendation.²¹ Indeed, the court might decide that for a clinician to abrogate their personal responsibility and instead delegate clinical decision-making to an AIS is conduct so specious that the claim could proceed on this ground. Consequently, if a clinician wishes to use AI technologies in treating patients and avoid breaching their duty of care, the cautious clinician would ideally be able to fully justify their decision-making independently of the AIS.

Ordinarily, the clinician’s relationship with the patient is clear, but using an AIS introduces the ‘third’ agent of the SDC. This complicates the identification of the entity which is the primary cause of harm. Neither the SDC, nor an advisory AIS which does not have physical contact with the patient would have been able to directly cause harm on their own; the clinician would have been the gateway to the AIS being able to harm via its recommendations to the clinician.

¹⁸ *Bolitho v City and Hackney Health Authority* [1998] AC 232

¹⁹ *Maynard v West Midlands Regional Health Authority* [1985] 1 All ER 635

²⁰ *As above*

Both *Simms v Simms and An NHS Trust* [2002] EWHC 2734 and *De Freitas v O’Brien* [1993] 4 Med LR 281, CA endorse minority opinions as being non-negligent.

²¹ Lord Browne-Wilkinson in *Bolitho v City and Hackney Health Authority* [1998] AC 232, 243 held that if ‘it can be demonstrated that the professional opinion is not capable of withstanding logical analysis, the judge is entitled to hold that the body of opinion is not reasonable or responsible’.

Hence, ‘but for’ the clinician’s conduct, the AIS’s harmful output would not have influenced the patient’s treatment at all.

The example of the use of IBM’s Watson provides a scenario to which causation may be discussed. Were a patient to come to harm due to inappropriate drug recommendations from an AIS being adopted, it might be possible to argue that ‘but for’ the AIS’s presence, the injury would not have occurred. A generalist clinician without specialist training in a specific illness (e.g. oncology), would not have been able to attempt to treat that illness; they would have been unable to choose or administer any specialist drugs as they would not have known which drugs could have been appropriate for the patient’s condition. Consider this hypothetical scenario. Suppose the clinician has no other colleague to refer the patient to, yet if timely treatment is not provided the patient could suffer. The clinician’s employing hospital has provided the AIS for this specific purpose. Under these conditions, a reasonable clinician may feel compelled to use that AI system. In this situation, ‘but for’ the presence of the AIS there would have been no attempt to treat the patient’s illness, thus no selection or administration of an inappropriate drug, but potentially averting more serious consequences from non-intervention.

In the absence of a valid defence, a claim made by a patient against a clinician who has used a defective AIS as part of their treatment plan may be irresistible. One could argue that the court would quite rightly find a clinician who follows an AIS’s recommendations without careful consideration of the consequences has acted with negligence. Yet it does not appear to be ‘fair, just and reasonable’²² for negligence liability to be extended only to clinicians, given the role of the SDC in a shared endeavour with clinicians in developing and deploying AIS. When an SDC claims that its product is both dispensing state of the art knowledge, but with the additional qualification that the clinician is to make the final decision of how to act on this

²² *Caparo Industries v Dickman* [1990] 2 AC 605

knowledge, it is not unreasonable for a clinician to consider rejecting the use of an AIS until it can be proven that its outputs can be safely relied upon. If an SDC is unable to provide adequate reassurance to clinicians, clinicians would be ill-advised to take the risk of using those systems.

It is the clinician's responsibility to ensure that they are familiar with any tools they use. This includes awareness of the risks of AISs. However, if the clinician cannot scrutinise the AIS because of the 'black box' character of the algorithm, they will be unable to take appropriate steps to mitigate these risks. This merits an assessment of the conduct of the SDC to determine whether they have materially contributed to the overall risk of harm to the patient.

IV. SOFTWARE DEVELOPER COMPANY'S NEGLIGENCE

The discussion of clinical negligence in Section III raised the possibility that the SDC might also share liability in negligence. For the purposes of a claim based on joint liability the conduct of the SDC must be analysed separately. *Donoghue v Stevenson*²³ provides a distinctive illustration of the duty of care manufacturers owe to the end users of their products. In *Donoghue*, a drink was served to consumers in opaque glass bottles, which rendered safety inspections futile once the bottles had been sealed at the factory. Extending this principle by analogy, it could be argued that the SDC relies on the clinician to intervene where an AIS recommends a course of action that is potentially unsafe. However, even where a clinician exercises professional judgement, it may be impossible to prove the extent to which the clinician has come to their decision independently, without any degree of influence from the AIS. Therefore, a court will need to assess whether the SDC has taken all appropriate measures to mitigate the risk of harm to patients where their AIS is deployed in clinical settings.

A. Duty of care

²³ [1932] AC 562

As the caselaw on negligence has developed, the duty of care has been extended ‘incrementally and by analogy’ with established duty situations.²⁴ Defendants cannot act with impunity, expecting that liability will rest with another party. The categories of relationship where a duty of care may be imposed can be expanded as novel situations arise. The *Caparo*²⁵ criteria provide guidance as to where the court may be persuaded to broaden the scope of negligence liability and determine if a duty of care exists. They query 1) whether the harm resulting from the defendant’s conduct reasonably foreseeable; 2) whether the relationship between the parties is sufficiently proximate in law; and 3) if it is ‘fair, just, and reasonable’ to impose a duty of care on the defendant. The following discussion sketches out the potential liability of the SDC towards the patient using this framework.

1. Foreseeability

SDCs cannot say that they did not foresee that patients would be affected by their system’s outputs given that their system was specifically designed to advise patient care. The relationship between the patient and the AIS is undiluted even if mediated by a clinician as the patient’s data is processed by the system directly. A skilled professional in possession of knowledge who wilfully ignores it is ‘prima facie negligent’ when it is foreseeable that an injury could occur.²⁶ This extends to technical professionals in the employ of the SDC.

The SDC could argue that intermittent episodes of malfunctioning are not unprecedented.²⁷ AISs operate according to observations latent in the data, rather than the experiential and applied knowledge base which clinicians possess. SDCs are not be expected to have medical

²⁴ *Robinson v Chief Constable of West Yorkshire Police* [2018] AC 736, [26]-[27]

²⁵ *Caparo Industries v Dickman* [1990] 2 AC 605

²⁶ D Rowland and JJ Rowland, ‘Competence and legal liability in the development of software for safety-related applications’ (1993) 2 *Information & Communications Technology Law* 229, 240

²⁷ T Petricek, ‘Miscomputation in software: learning to live with errors’ (2017) 1(2) *The Art, Science and Engineering of Programming* <<https://programming-journal.org/2017/1/14/>> accessed 8 March 2020

expertise, but a negligence claim might not be unreasonable if an SDC has released software for use in clinical environments.

An alternative foreseeability scenario also exists. Should the AIS have a high frequency of accurate outputs desirable for clinical decision-making, the clinician may find their attention wanes when monitoring the effect of the system's recommendations as applied to patients. This phenomenon can be described as an 'atrophy of vigilance'.²⁸ When the device works consistently without issue, the user might begin to trust it even when they know that they should not, but this is in conflict with the clinician's duty of care which would compel them to pay attention when using any kind of tool.

If atrophy of vigilance can result in death or personal injury in the above safety critical situations, it might be considered a foreseeable consequence for AISs which provide rapid solutions and aim to lessen the cognitive burden for system users. Trusting an AIS which is usually reliable is a foreseeable consequence in the lifespan of an AIS; thus, it could be posited that there is an open the door for courts to find that SDCs owe a duty of care if harm eventuates due to a clinician's foreseeable loss of attention whilst using their system. Were the courts to agree, a positive obligation from SDCs could be required to take human factors into account and design the system to ensure safety in the high-pressured clinical environment. Absence of such holistic design features could be taken as evidence of a negligent omission.

2. Proximity

To avoid liability claims, the SDC would attempt to position the clinician as the primary guardian of safety. Proximity, as understood in the tort of negligence, need not be solely geographical, instead it describes

²⁸ WR Freudenberg, 'Nothing recedes like success? Risk analysis and the organizational amplification of risks' (1992) 3(1) *Risk: Issues in Health and Safety* 1

‘such close and direct relations that the act complained of directly affects a person whom the person alleged to be bound to take care would know would be directly affected by his careless act’.²⁹

Within the clinical environment, this duty of care towards patients is held by both medical and non-medical staff as found in *Darnley*.³⁰

It is conceivable that a similar relationship might be considered to exist between an SDC and a patient if the SDC provided their system to be used specifically to aid patient care. The SDC’s position might be considered analogous to other professional roles within the clinical environment, for example radiologists or haematologists. These specialised clinicians can advise their colleagues on an interpretation of a medical image or on the appropriateness of administering a unit of blood product to a specific patient; and they frequently do this without ever having met the patient themselves. Specialised clinicians who are not directly at a patient’s bedside are expected to follow the same code of professional conduct as their fellow colleagues.³¹ In law, specialists are subject to the same duty of care as generalists, though the expected standard of care may differ.

This matrix can be applied to the SDC which develops an AIS which accepts an input, processes it according to a specialised algorithm, and then generates a recommendation which purports to embody clinical expertise. However, the SDC is a third party detached from the bedside. Thus, an AIS may make recommendations which prompt a user’s actions, but this does not lessen a user’s duty to take care that their actions shall not harm another. Here we see

²⁹ *Donoghue v Stevenson* [1932] AC 562, 581

³⁰ *Darnley v Croydon Health Services NHS Trust* [2018] UKSC 50

³¹ General Medical Council (2013) *Good medical practice: General Medical Council*. <<https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/good-medical-practice>> accessed 8 March 2020
Health & Care Professions Council (2016) *Standards of Conduct, Performance and Ethics*. <<https://www.hcpc-uk.org/publications/standards/index.asp?id=38>> accessed 8 March 2020
Nursing and Midwifery Council (2018) *The Code for Nurses and Midwives*. London: Nursing and Midwifery Council. <<https://www.nmc.org.uk/standards/code/read-the-code-online/>> accessed 8 March 2020

that the participation of the SDC complicates an otherwise straightforward assessment of where the duty of care lies.

3. 'Fair, just and reasonable'

The third component of the *Caparo* guidelines is more challenging to apply, not least because it is seen as ancillary to the requirement of proximity and that it encapsulates policy considerations that may more usually be the province of Parliament.³²

Firstly, there is no general duty of care to prevent damage being inflicted by third parties.³³ Given that the AIS is unable to dispense treatments without the cooperation of a clinician, this is a reason for denying the extension of the duty of care to encompass SDCs. But the AIS has been designed specifically to influence the actions of clinicians to directly affect the health status of the patient. The general rule is that there is no duty of care for persons to prevent harm being caused by third parties, yet a 'special relationship' may be possible, which would override the conduct of others; as in, for example, the recent case of *ABC v St George's Healthcare NHS Trust*.³⁴ Here, the Court of Appeal found that a claim based on a 'special relationship' between an NHS trust and a patient's unborn grandchild was sufficient to be arguable at trial³⁵ and was recently confirmed on the facts in the High Court.³⁶ Thus, it is not inconceivable that the courts might consider that there are grounds to argue that a 'special relationship' exists between the patient and the SDC, even though there is no case law to illustrate a duty of care between the SDC and the patient at present. If the system has been

³² *Alcock v Chief Constable of South Yorkshire* [1992] 1 AC 310, 410; *Michael v Chief Constable of South Wales* [2015] AC 1732, [156]-[158]

³³ *Smith v Littlewoods Organisation Ltd* [1987] AC 241

³⁴ [2017] EWCA Civ 336; 160 BMLR 19

³⁵ J Mackenzie (2017) *The Medical Duty of Care to a Third Party*

<<https://www.anthonygold.co.uk/latest/blog/medical-duty-care-third-party/>>accessed 8 March 2020

³⁶ [2020] EWHC 455.

deployed with that specific purpose, it could be argued that it is unreasonable that the SDC's legal responsibility for the effect of the system is negated by the clinician using it.

Secondly, an SDC may argue that an AIS is no substitute for skilled, clinical decision-making and that the law only imposes a standard of care commensurate with the level of specialism which the defendant holds themselves out as possessing.³⁷ . If a SDC wished to transfer legal responsibility for using the AIS to the clinical user, they would therefore present it as merely 'assisting' clinicians in their decision-making, but would not claim that it is of the standard to substitute a clinician in the speciality that the system advises in. Based on this premise, it might be argued that it is unfair and unreasonable to assign negligence liability to the SDC which created and supplied that system.

A clinician would, however, be unlikely to consider using an AIS they believe would make recommendations which are inferior to their own calculations; to do so would be illogical, counter-productive and would expose a patient to needless risk. In the example of IBM's Watson for Oncology, the system is portrayed by the SDC as possessing the expertise of 'leading oncologists'.³⁸ The reported use of this system by UB Songdo hospital³⁹ is concerning if they had been informed that Watson makes recommendations rather than clinical decisions,⁴⁰ thus resulting in the risk of system use being borne by its clinical user. UB Songdo Hospital must have been convinced that Watson was sufficiently sophisticated to perform that role; in

³⁷See *Philips v William Whiteley Ltd* [1938] 1 All ER 566 where the court held that a jeweller did not claim to be of the same standard as a surgeon, and that the appropriate standard of care was of a reasonable jeweller undertaking ear piercing and not the reasonable surgeon.

³⁸ IBM (undated) *IBM Watson for Oncology: What Watson for Oncology can do for your organization* <<https://www.ibm.com/products/clinical-decision-support-oncology>> accessed 8 March 2020

³⁹ C Ross and I Swetlitz, (2017) 'IBM pitched Watson as a revolution in cancer care. It's nowhere close'. *STAT News*. <<https://www.statnews.com/2017/09/05/watson-ibm-cancer/>> accessed 8 March 2020

⁴⁰ M Hengstler, E Enkel and S Duelli, 'Applied artificial intelligence and trust – the case of autonomous vehicles and medical assistance devices' (2016) 105 *Technological Forecasting and Social Change* 105, 115

the absence of universally accepted standards, individual institutions must assess each AIS's adequacy on a system-by-system basis.

A preference for extending the duty of care to the SDC might exist if there is sufficient foreseeability of harm and a convincing relationship of proximity between the SDC and the patient. This could be too challenging for the courts to accept as the clinician is acting at the bedside whereas the SDC is not. Yet, the SDC is acting for the benefit of the patient notwithstanding that the clinician is the intended recipient of the AIS's recommendations and there is an interesting analogy here with the case of *Smith v Eric S Bush*.⁴¹ The defendant surveyor was instructed by a building society to report on a property being purchased by the claimant. In addition to the surveyor owing a duty of care to the building society to carry out the inspection with due skill and care, the surveyor owed a duty to the third party purchaser on the basis that the defendant would know there was an 'overwhelming probability' that the purchaser would also rely on that report. Similarly, there is an 'overwhelming probability' that if the recommendation given to the clinician is negligent, its impact will be felt by the patient. This analogy reinforces the argument for extending the duty of care to the remote patient.⁴² Similarly, if the AIS is a 'black box', meaning that the clinician cannot interrogate its reasoning, it might be 'fair, just and reasonable' that the SDC should owe a duty of care to the patient too, independently of any negligent conduct on the part of the clinician.

The *Caparo* principles were not intended to cover all future scenarios, but to guide the court's consideration of novel fact situations. We argue in favour of imposing a duty on the SDC in addition to any claims against the clinician so that risks are appropriately managed and contained.

⁴¹ [1990] UKHL 1.

⁴² [1990] UKHL at [26], *per* Lord Griffiths.

B. Breach of duty

As previously mentioned, clinical conduct is not considered negligent if it satisfies the *Bolam-Bolitho* calibration of the standard of care. If a duty of care is found to exist for the SDC, their system would need to be in accordance with a responsible body of opinion and the expert evidence supporting the defendant's conduct or decision must 'withstand logical analysis'. We consider how this calibration of the standard of care could apply to the SDC.

The AIS might advise contrary to the expectations of the clinician, but that might not necessarily be negligent as long as the system's recommendations are safe and therapeutic for the patient.⁴³ In safety critical areas it is desirable that an SDC strives to follow, and even surpass, the relevant standards for medical devices.⁴⁴ Healthcare is recognised as a safety critical area; any incorrect advice from an AIS has the potential to harm the target patient.

An SDC could demonstrate discharging their duty of care through observation of the standards and codes of practice relevant to their profession.⁴⁵ The notion of observation of standards is supported in comparative case law from New Zealand. In *Bevan Investments v Blackhall & Struthers*⁴⁶ the defendant engineer had achieved the expected standard, and so it was held that upon rational analysis the court could conclude that no negligence had occurred. Codes of conduct and professional standards are in place for clinicians,⁴⁷ however technical standards

⁴³ *Luxmore-May v Messenger May Baverstock* [1990] 1 All ER 1067

⁴⁴ D Rowland and JJ Rowland, 'Competence and legal liability in the development of software for safety-related applications' (1993) 2 *Information & Communications Technology Law* 229

⁴⁵ D Rowland and JJ Rowland, 'Competence and legal liability in the development of software for safety-related applications' (1993) 2 *Information & Communications Technology Law* 229, 236

⁴⁶ (1979) 11 BLR 78

⁴⁷ General Medical Council (2013) *Good medical practice: General Medical Council*. <<https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/good-medical-practice>> accessed 8 March 2020
Health & Care Professions Council (2016) *Standards of Conduct, Performance and Ethics*. <<https://www.hcpc-uk.org/publications/standards/index.asp?id=38>> accessed 8 March 2020
Nursing and Midwifery Council (2018) *The Code for Nurses and Midwives*. London: Nursing and Midwifery Council. <<https://www.nmc.org.uk/standards/code/read-the-code-online/>> accessed 8 March 2020
National Institute for Health and Care Excellence (2020) <<https://www.nice.org.uk/about>> accessed 8 March 2020.

are not yet sufficiently advanced in the domain of AI assisted medicine to provide definitive guidance to SDCs.

Generalised codes of conduct exist in many forms, but without teeth or unification.⁴⁸ Clinical practice standards need to be sector-specific, or even health condition-specific before medical experts and the public could be reliably assured of the safety of AI devices and algorithms. UK clinicians are aided by organisations such as the National Institute of Health and Care Excellence (NICE) which publishes evidence-based guidance on clinical conditions and treatment pathways.⁴⁹ However, NICE has not evaluated AIS for use in treatment pathways. Instead, NHSX, a unit dedicated to digital transformation of the NHS,⁵⁰ is working on standard setting in this area. This is supported by the release of NICE's Evidence Standards Framework for Digital Health Technologies.⁵¹ Whilst we do not have a body of knowledge from which the courts may draw comparisons at this moment in time, initiatives such as this will begin to build that knowledge base and the standards of conduct which SDCs should adhere to for the courts to draw upon in the future. Although the standard of care in negligence and the standards set out in the various sources of published guidance are not *necessarily* the same, courts often rely heavily on professional guidance to inform the standard the law should apply. For example, in

⁴⁸ Access Now, *The Toronto declaration: protecting the rights to equality and non-discrimination in machine learning systems* (16 May 2018) <www.accessnow.org/the-toronto-declaration-protecting-the-rights-to-equality-and-non-discrimination-in-machine-learning-systems> accessed 8 March 2020; Association for Computing Machinery, *ACM code of ethics and professional conduct* (2018) <www.acm.org/code-of-ethics> accessed 8 March 2020; Future of Life Institute, *Asilomar principles* (2017) <<https://futureoflife.org/ai-principles>> accessed 3 June 2019; Institute of Electrical and Electronics Engineers, *Ethically aligned design* (2019) <<https://ethicsinaction.ieee.org>> accessed 8 March 2020; The Public Voice, *Universal guidelines for artificial intelligence* (2018) <<https://thepublicvoice.org/ai-universal-guidelines/memo>> accessed 8 March 2020.

⁴⁹ National Institute of Health and Care Excellence <www.nice.org.uk/guidance> accessed 8 March 2020

⁵⁰ Department of Health and Social Care, 'NHSX: new joint organisation for digital, data and technology' (GOV.UK, 19 February 2019) <<https://www.gov.uk/government/news/nhsx-new-joint-organisation-for-digital-data-and-technology>> accessed 8 March 2020

⁵¹ National Institute for Health and Care Excellence, *Evidence Standards Framework for Digital Health Technologies* (2019) <<https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>> accessed 8 March 2020

*Montgomery v Lanarkshire Health Board*⁵² the Supreme Court expressed doctors' duties to inform patients as closely aligned with General Medical Council (GMC) guidance on the matter.

If the AIS is thought to be so risky that its outputs need to be verified by a clinician, the court might find it hard to accept that a system's outputs are effective substitutes for the professional standards of clinical staff (and that it thereby does not satisfy the *Bolam-Bolitho* standard of care). AIS outputs are not the same as the evidence-based medicine upon which modern clinical practice is grounded. However, the clinician may not have the skills to appraise the AIS which has been offered to them. Additionally, there is no peer reviewed evidence base which the clinician may draw upon to ensure that using the AIS would result in an improvement in care. These two issues create a problem of non-translatability from the body of knowledge offered by the AIS to the clinical environment. When presenting an AIS for appraisal and use by clinicians, a SDC would need to be circumspect in their conduct to avoid misrepresentation of their product as equivalent to the work of 'the ordinary skilled man exercising and professing to have that special skill'.⁵³ Advertising exaggerations regarding AIS's may shape liability under the Consumer Protection Act⁵⁴ and give rise to claims under contract law for misrepresentation, but the crux of the matter is that the conduct of SDCs as interdisciplinary specialists could be used as evidence that the SDCs ought to have known that defects in their system would be highly likely to inflict harm, putting them in breach of their duty of care for not taking active steps to avoid such harm.

C. Causation

This paper has already touched upon causation in the earlier discussion of factual causation in section III. Broadly speaking, the two major views are that either a) 'but for' the malfunctioning

⁵² [2015] UKSC 11

⁵³ *Bolam v Friern Barnet Hospital Management Committee* [1957] 2 All ER 118, 121

⁵⁴ Consumer Protection Act 1987, s 3(2)

AIS the clinician would not have been prompted to apply the harmful recommendation to the patient, or b) that a clinician openly acknowledging that they do not have the skills of a specialist clinician does not relieve them of a negligence claim should harm eventuate when they have used an AIS which makes recommendations in that specialist area. However, if the court holds that the SDC shares responsibility with the clinician for the patient's harm, the court must consider how multiple causes may be related.

When multiple factors and/or multiple tortfeasors are determined as having caused injury, the courts use different methods to determine the cause responsible for the injury. The tortfeasors' 'independency' approach finds that each factor had caused a single damage independently of other factors.⁵⁵ By contrast, the tortfeasors' 'separate impact' approach differs from tortfeasor's independency as multiple injuries are attributed to separate tortfeasors.⁵⁶ Tortfeasors' contribution exists when multiple tortfeasors all acted and the sum of the acts are treated as creating the damage.⁵⁷ The use of these three approaches can make it challenging to predict the outcome of a claim.⁵⁸

Suppose the clinician and the AIS independently make identical errors. Specifically, if a defective AIS produced an erroneous output which could harm the patient if followed, is the effect of this error eliminated by the clinician's independent conduct? In *Performance Cars v Abraham*⁵⁹ it was held that the first act may obliterate the second. One may be able to say that both actions were tortious and conclude that the effect of the first act (the negligent development of the AIS) continues in spite of subsequent negligence.⁶⁰ The SDC's negligence would take priority in this instance, because reference to the AIS would precede the clinician's

⁵⁵ E Noee, M Noee, and A Mehrpouyan. (2016) 'Attribution of Liability among Multiple Tortfeasors under Negligence Law: Causation in Iran and England.' *Journal of Politics and Law*. 9 (7) p.219-229 p.221

⁵⁶ Ibid. p.224

⁵⁷ Ibid. p.223

⁵⁸ Ibid. p.221

⁵⁹ [1962] QB 33

⁶⁰ *Performance Cars Ltd v Abraham* [1962] 1 QB 33; *Heil v Rankin* [2001] PIQR Q3

conduct. However, this might unduly relieve clinicians from acting in the patient's best interests and raise a moral hazard, encouraging reckless behaviour. As such, the 'consecutive cause' approach may not provide the best means for balancing the responsibility of the clinical and technical parties in ensuring that the AIS is developed and used in a way that guarantees patient safety. Furthermore, it is questionable whether the *Performance Cars* doctrine is applicable here as the 'conduct' of the SDC, as expressed through the AIS tool, and that of the clinician are not truly independent of each other.

Where the 'but for' approach to causation fails, the patient may be able to satisfy the causation element of a claim by proving that the negligence 'materially contributed to the damage', and if successful, the defendant may be liable in full for the whole of the damage, notwithstanding that they have only been proved to contribute to it.⁶¹ However, where the damage suffered is treated as 'indivisible' (as seems to be the case with many cancers), the parties might argue that liability on the basis of being responsible for a 'material increase of risk' is the more appropriate claim.⁶² The jury is still out however on whether this more liberal approach to proving causation will ever be applied outside the so called 'toxic torts' or industrial diseases cases.⁶³ If arguable, the material increase of the risk approach to causation might be applied where both the clinician and the AIS independently make the same error but as each confirms the other, it is impossible to prove which entity contributed more to the harm suffered by the patient. But, both aforementioned arguments depend on the SDC's AIS being accepted as an authoritative source. If the recommendations given by an AIS is not accepted as such, it could

⁶¹ See *Bonnington Castings v Wardlaw* [1956] AC 613, *Bailey v Ministry of Defence* [2008] EWCA Civ 1052 and *John v Central Manchester and Manchester Children's Hospital University Hospitals NHS Trust* [2016] EWHC 407.

⁶² *Fairchild v Glenhaven Funeral Services Ltd* [2003] 1 AC 32

⁶³ See e.g. *John v Central Manchester and Manchester Children's Hospital University Hospitals NHS Trust* [2016] EWHC 407.

be argued that the clinician should have made their own independent check on the appropriateness of an AIS's recommendations.

If a court were to approve of the 'material increase of risk' approach to proving causation, it is necessary to outline the limitations on obtaining compensation that this entails. According to *Barker*,⁶⁴ liability for a 'material increase of risk' is to be apportioned severally, meaning that a patient claimant or a clinician co-defendant may not be able to claim the appropriate contribution to damages from the SDC. Several liability is recommended by the European Commission whose report directly considered liability for AISs.⁶⁵ Yet, under a several liability regime, the claimant must succeed in their claim against each defendant separately in order to receive compensation in full for the negligently inflicted injury, which adds additional unnecessary costs and minimises the prospect of obtaining full compensation for the harm suffered.⁶⁶ This could discourage claimants from pursuing justice from an SDC and mean that the incentive for SDCs to fulfil their duty of care is weakened. In Section VI below, we argue for a return of joint liability to ensure that the patient has access to a timely legal remedy.

In Section III it was shown that the AIS is unable to inflict harm without the conduct of the clinician as an intermediary. It would therefore be advantageous to the SDC to argue that the conduct of the clinician is a *novus actus interveniens*. Actions of third parties do not generally break the chain of causation unless the intervening conduct is so outrageously negligent that it would not be fair for the initial defendant to continue to carry responsibility for those later acts.⁶⁷ Suppose the clinician had failed to recognise that a treatment recommendation was inappropriate to the patient, despite a reasonable expectation that they ought to have known

⁶⁴ *Barker v Corus UK Ltd* [2006] 2 AC 572

⁶⁵ Expert Group on Liability and New Technologies – New Technologies Formation. (2019) *Liability for Artificial Intelligence and other emerging digital technologies*. European Union: European Commission <<https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608>> accessed 8 March 2020

⁶⁶ *Barker v Corus UK Ltd* [2006] 2 AC 572

⁶⁷ *Spencer v Wincanton* [2009] EWCA Civ 1404

such a treatment could harm the patient. If they then acted solely on the basis of the defective AIS recommendation, the clinician may well find themselves causally responsible. In this case, the court may be minded to hold that these are two separate instances of negligent acts. Instead of the *Performance Cars* approach where the harm caused by the initial act is deemed to continue, here the situation is reversed. Under the *novus actus interveniens* doctrine, any ‘extraordinary’ latter conduct of the clinician could obliterate the defects latent in the AIS, which may remain undiscovered. If the spirit of *Bolam* is that the clinician would need to have achieved the standards of other professionals, then the clinician may have failed to achieve this standard if they had failed to identify that an AI’s recommendation was inappropriate for a specific patient.

Does the clinician professing to their patient that they do not have the skills of a specialist clinician relieve them of the usual standard of care? Likely not when considering the objective standard of care as illustrated by *Nettleship v Weston*;⁶⁸ if ‘the certainty of a general standard is preferable to the vagaries of a fluctuating standard’⁶⁹ then the standard expected of a clinician offering specialist treatment for a specific condition ought be determined of an acceptable standard as per a specialist clinician’s practice, rather than of a non-specialist clinician offering specialist treatment of which they are not qualified. This principle was confirmed in the recent case of *FB v Princess Alexandra Hospital NHS Trust*⁷⁰ where it was found that, when considering liability, the standard of a doctor’s care is not to be adjusted to take into account their limited experience.

In the absence of a definitive pronouncement by the court on the causation issues raised in this paper regarding the presence of the AIS, it is impossible to accurately predict how far causation

⁶⁸ [1971] 2 QB 691

⁶⁹ *Ibid.* at 707

⁷⁰ [2017] EWCA Civ 334

principles will be central to the future development of the law. Nevertheless, we have shown that the common law is not closed on these matters, meaning that there is a real possibility that SDCs may be liable in negligence for defective AISs in clinical settings.

D. Using volenti as a defence

The characterisation of IBM Watson for Oncology as a recommendation engine suggests that the technical community may adopt a strategy of excluding their liability in negligence.⁷¹ The injured patient would be the claimant, so are not party to a contract between the SDC and the clinician. Thus, despite their pronouncements that the clinician retains ultimate responsibility for patient care, in law this is not the SDC's choice to make. A defendant cannot exclude liability for negligently caused death or personal injury,⁷² so this tactic might prove to be an inadequate defence in a negligence claim brought by a patient.

Thus, a defendant SDC's option for a full defence might be limited to arguing *volenti non fit injuria*. In medical contexts, this principle contains the following elements: consent must be given voluntarily by the patient⁷³ and they must be of sound mind.⁷⁴ The patient is also entitled to be provided with information about all the relevant factors so that they can formulate their decision.⁷⁵

Questions could be raised about whether it is possible for the patient to voluntarily agree to the use of a 'black box' AIS, where its processes are largely unknowable and inscrutable in precise detail. It is reasonable that a patient would want to know if a system they were about to use might produce faulty outputs which might harm them before consenting to its use. A prudent patient may wish to demand that the SDCs quantify the risks before engaging with their

⁷¹ M Hengstler, E Enkel, and S Duelli, 'Applied artificial intelligence and trust – the case of autonomous vehicles and medical assistance devices' (2016) 105 *Technological Forecasting and Social Change* 105

⁷² Unfair Contract Terms Act 1977 s 2(1)

⁷³ *Smith v Charles Baker & Sons* [1891] AC 325

⁷⁴ *Kirkham v Chief Constable of Greater Manchester* [1990] 2 QB 283

⁷⁵ *Chester v Afshar* [2005] 1 AC 134

products. With this knowledge they may choose to not permit the use of an AIS in their care. On the other hand, if the patient did choose to permit use of the AIS knowing of the risk of potentially harmful recommendations being produced, they would be wise to consider a low threshold to seeking verification by a clinician prior to accepting the treatment which the AIS had recommended. *Volenti* in any case is a very unlikely defence in a medical negligence action; it is only rarely successful and in very limited circumstances.⁷⁶

V. RISK POOLING

The present state of the law does not account for the way that SDCs might perpetrate harms from a distance. In conventional negligence liability, the principles of proximity (with regard to the duty of care) and remoteness (in respect of the damage caused) are used to delimit the legal responsibility for acts or omissions to the parties that are most directly connected to the harm. This arises out of the notion that ‘fault’ ought to be the basis of liability;⁷⁷ the broader the definition of liability, the greater the likelihood that ‘fault’ is applied to parties whom society might otherwise perceive as morally innocent, or whose conduct is justifiable and excusable.

Another basis for awarding damages is that it provides a deterrence signal to potential tortfeasors.⁷⁸ Ultimately, claimants would rather that their injuries had never occurred; the threat of being forced to pay financial compensation is one way to encourage parties to consider their actions carefully. Closely related to this is the idea that liability promotes the efficient ‘internalisation’ of costs in risky activities.⁷⁹ Therefore, the defendant is generally the party

⁷⁶ See J. McHale and J. Laing, (eds) *Principles of Medical Law* (OUP, 2010) at 6.73.

⁷⁷ P Cane and J Goudkamp, (2018) *Atiyah's Accidents, Compensation and the Law*. Cambridge: Cambridge University Press

⁷⁸ D Schwarcz and P Siegelman, (2015) *Research Handbook on the Economics of Insurance Law*. Cheltenham, UK: Edward Elgar p.96

⁷⁹ M Faure and RA Partain, (2017) *Carbon capture and storage : efficient legal policies for risk governance and compensation*. London: The MIT Press p.103

who could have avoided the harm at the least cost. Hence, compensation is a redistributive obligation that seeks to restore economic equilibrium amongst participants.

It is this latter view that reflects the position of the three parties introduced in Section II. The clinician has a clear obligation to act in the best interests of their patient. The SDCs ought to design their system so that is ‘fail safe’; minimising the risk of harm when defects occur and forestalling the ‘atrophy of vigilance’ phenomenon. Likewise, the patient could have taken steps to avoid harm by interrogating the safety aspects of the AIS but has been prevented from doing so due to the ‘black box’ character of these systems (thus their consent is vitiated).

The patient is clearly deserving of compensation if they have suffered unnecessary harm, so the question becomes one of how the patient claimant’s loss is best distributed and rectified. The present article argues that liability for damages ought to be shared among the stakeholders in medical AI devices and algorithms, involving both the technical and medical teams as they jointly contribute to the overall risk of harm. Yet, there is also a concurrent tension inherent in the desire to innovate and provide an environment where beneficial technologies can be tested and deployed rapidly in front-line care.

The UK government proposes that liability insurance may help to balance the risk between actors and provide clear accountability among participants in sensitive sectors.⁸⁰ If an actor can envisage a duty of care arising from their actions, insurance can help to defray some of the costs of engaging in risky activity. Indeed, in some instances, purchasing insurance to cover the minimum expected liability is compulsory.⁸¹ But in order to insure against harms, there needs to be a clear model of how restitution for that loss is to be implemented.

⁸⁰ Government Office for Science, *Artificial intelligence: opportunities and implications for the future of decision making* (GS/16/19, 2016) 15
<https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf> accessed 8 March 2020

⁸¹ e.g. Road Traffic Act 1988, ss 143 and 145

Allain, writing from a US perspective, suggests that the ‘enterprise liability’ approach may satisfy the interests of all parties.⁸² Enterprise liability is broadly identified by Klemme as the imperative that ‘losses to society created or caused by an enterprise or, more simply, by an activity, ought to be borne by that enterprise or activity’.⁸³ The fault of each actor within an ‘enterprise’ might well be covered by either a singular insurance paid for by both the SDC and the clinician or by separate policies which reduce the impact on each individual in the event of a claim.⁸⁴

Allain proposed that the user of devices such as IBM Watson be indemnified with insurance which combines aspects of product liability and vicarious liability as well as medical malpractice and allows the spread of fault between the clinicians using the system, the SDCs who have developed the system and the hospital where the system is being used. The enterprise liability model reduces the burden on claimants as the court will not need to analyse each actor’s role in the claimant’s misfortune, but instead look at the actions of the team *in toto*. In this way, ‘insurance acts to better spread the risk of loss throughout society reducing the economic impact of each individual judgment’.⁸⁵

Allain proposes that restitution to the claimant would be shared equally between the SDC who has supplied the system and the clinician or hospital who has adopted the system.⁸⁶ She argues that if stakeholders equally share the burden of loss, the risk of loss is shared across all actors resulting in reduced economic disincentives. This would encourage SDCs to ensure that their system is as accurate as possible, and hospital management teams can be reassured that they

⁸² JS Allain, ‘From Jeopardy! to jaundice: the medical liability implications of Dr Watson and other artificial intelligence systems’ (2013) 73 *Louisiana Law Review* 1049

⁸³ HC Klemme, ‘The Enterprise Liability Theory of Torts’ (1975–6) 47 *University of Colorado Law Review* 153, 158.

⁸⁴ JS Allain, ‘From Jeopardy! to jaundice: the medical liability implications of Dr Watson and other artificial intelligence systems’ (2013) 73 *Louisiana Law Review* 1049, 1077

⁸⁵ JS Allain, ‘From Jeopardy! to jaundice: the medical liability implications of Dr Watson and other artificial intelligence systems’ (2013) 73 *Louisiana Law Review* 1049, 1077

⁸⁶ JS Allain, ‘From Jeopardy! to jaundice: the medical liability implications of Dr Watson and other artificial intelligence systems’ (2013) 73 *Louisiana Law Review* 1049, 1074

will not be left to shoulder the full cost without appropriate contributions or reimbursement from their technical partners.

However, it is debatably not enterprise liability if the economic impact of a claim is assigned to an insurer rather than directly impacting the actors who caused the harm. Loss-spreading using insurance allows the shifting of liability to almost a community-wide basis. This paper identifies that loss-spreading and enterprise liability are distinct, and the scenario relies upon actors purchasing insurance which covers their activities.

Insurance is beneficial if society accepts that the introduction of AIS in clinical decision-making is, on balance, beneficial to the public and the adoption of this technology is deemed as collectively in society's interest. If the SDC and the clinician must obtain insurance for use of an AIS to be permissible in the clinical decision-making context, the cost of paying for that insurance shall fall upon those who pay for the healthcare provided. Therefore, rather than considering the SDC and the clinician as paying for the required insurance, it is in fact the patient or those who pay for the patient's care who are burdened with that cost; the community truly does carry the cost of liability when seen this way.

Enterprise liability may lead to parties other than the insurer being liable. For example, if the clinician uses a recommendation which leads to patient injury and the insurance policy's wording does not cover this tort, the insurance might not benefit the injured patient; thus, negating the enterprise's intention of obtaining insurance. As a solution, this paper identifies the use of 'risk pooling' to reflect stakeholders' collective acceptance and preparation for the potential for harm to arise when AISs are used in clinical decision-making.

Song defines risk pooling as:

'...if X performs an action which imposes an unreasonable risk of harm on Y, then X is liable to Y, and therefore obliged to make an ex ante compensation into a social pool

*that is roughly equivalent to the cost of expected harm (i.e., the probability of actual harm multiplied by the amount of the cost incurred by the harm).*⁸⁷

Merkin and Steele speak of insurance operating under an ‘actuarial model’ to spread risk across a discrete pool.⁸⁸ A risk pool in this paper’s context could be a single insurance policy which charges each participant as per their own risk. Merkin and Steele describe the principle of stakeholders receiving what they have paid in⁸⁹ which would be fair to other members of the risk pool. Yet, they warn that although risk pools may be well defined, they are rarely homogenous and that there is no requirement for premiums to be allocated as per the class of risk,⁹⁰ and so premiums could be unfairly costed between stakeholders. For example, the clinician is closer to the patient and is in the position of being the final decision-maker in clinical decision-making. Therefore, they hold a larger proportion of immediate responsibility than the SDC. If the clinician has not used the system in the way intended, it is not just or reasonable for the SDC to subsidise the clinician’s wrongs. Nevertheless, if the clinician has used the AIS appropriately and was unable to detect a system error, the SDC ought to make a contribution commensurate with their involvement in creating the risk of harm in the first place.

Additionally, there is the problem of ‘moral hazard’: just because an act is insured does not mean that it is free of responsibility.⁹¹ Thus, if an AIS is insured by all concerned stakeholders and there is still a potential risk of injuring a patient through using it, the provision of insurance cover does not morally excuse that the AIS had been deployed when foreseeable harm could take place. If it is preferable that AISs are tools and that human actors must not pass responsibility for harms to AISs, intentionally shifting responsibility for harms from human

⁸⁷ F Song, (2018) ‘Regarding a Risk-Pooling System of compensation.’ *Ratio*. 32(1)139-149 p.140

⁸⁸ R Merkin, and J Steele, (2013) *Insurance and the law of obligations*. Oxford, United Kingdom : Oxford University Press p.139

⁸⁹ Ibid. p.139

⁹⁰ Ibid. p.141

⁹¹ R Merkin, and J Steele, (2013) *Insurance and the law of obligations*. Oxford, United Kingdom: Oxford University Press p.141

actors to an insurance policy could be seen as similarly responsibility-avoidant behaviour. But whilst the use of AI in clinical decision-making is currently novel, the use of insurance to offset non-intentional risk has long been accepted by society. If society accepts the use and risks of, using AI in clinical decision-making, it is conceivable that risk pooling would be accepted here too for non-intentional harms.

A. Risk pooling in England and Wales: proposals for reform

There are no publicly discussed examples of application of insurance to AISs used in clinical decision-making, but the UK's preparation for self-driving cars might be considered analogous. The Automated and Electric Vehicles (AEV) Act 2018 effectively enacts the proposal for risk pooling as the basis for a compulsory insurance scheme in the context of self-driving vehicles. This legislation requires that, in the event of an accident due to an automated car 'driving itself', the insurer is liable to pay out on any meritorious claim.⁹² The injured party has priority ahead of all other actors within the scenario and compensation is made available to the claimant without them facing the prospect of a lengthy and expensive court battle. The insurer may then later seek recovery from other parties whose conduct contributed to the accident, for example the vehicle manufacturer or other road users.⁹³

The scenario of an AIS being used for clinical decision-making would be comparable to what is described as a highly automated vehicle in the AEV Act. When a highly automated vehicle is used, a human remains in the loop; the Law Commission in this context advises that the user-in-charge of the vehicle must remain either in the vehicle or within direct sight, they also retain responsibility to ensure roadworthiness and insurance.⁹⁴ Thus, it may be envisaged that the user

⁹² Automated and Electric Vehicles Act 2018 s 2(1)

⁹³ Automated and Electric Vehicles Act 2018 s 5

⁹⁴ Law Commission Consultation Paper No 245 and Scottish Law Commission Discussion Paper No 169. 16 October 2019. *Automated Vehicles: Consultation Paper 2 on Passenger Services and Public Transport: A joint consultation paper* < <https://www.lawcom.gov.uk/project/automated-vehicles/> > accessed 8 March 2020 p.2 and p.3

of an AIS being used for clinical decision-making would need to be present to be considered the user and they should expect to be similarly held responsible to ensure suitability and appropriateness of the AIS they use. Because a human remains in charge of choosing to use the outputs from an AIS, an AEV-style Act could be of no additional benefit to the clinician as it would specify them as being in charge if they were the final user. Thus, if they acted negligently, a user could expect the insurer to attempt to recover from them the relevant proportion of damages awarded to an injured party.⁹⁵

The AEV Act is a proactive law which seeks to lessen uncertainty when considering the adoption of future technologies whilst risking regulatory disconnect when technology is later introduced.⁹⁶ This paper has noted that there is a regulatory disconnect between the clinician, the SDC and the patient which leaves the clinician vulnerable to being burdened by potential negligence claims; there is an opportunity to proactively recognise this now so that shared responsibility for the use of an AIS in the clinical area could be realised via risk pooling in the future.

The court may engage in detailed consideration of the legal issues as they apply to the defendants, for instance establishing whether the duty of care exists or dealing with problems of causation. Thus, additional distress is avoided by the injured party and they can commence their recovery journey sooner. A risk pooling insurance scheme would cover harms when they arise, but unlike Allain's proposal, the defendants only contribute in accordance with the extent of their liability.

1. Insurance schemes for clinical malpractice should include coverage for AI-related damage

⁹⁵ M Channon, (2019) Automated and Electric Vehicles Act 2018: An Evaluation in light of Proactive Law and Regulatory Disconnect *European Journal of Law and Technology* 10(2)

⁹⁶ Ibid.

Clinicians have a professional obligation to carry adequate insurance against injuries arising from their tortious conduct.⁹⁷ They may additionally be subject to vicarious liability rules, whereby the employer is deemed to adopt the liability of the conduct of its employees. NHS bodies operate under a regime of self-insurance which is administered through NHS Resolution schemes.⁹⁸ Private healthcare institutions may optionally take out public liability insurance which may supplement or replace the insurance coverage of the individual clinicians who work for them. SDCs can purchase public liability insurance which can indemnify their liability in negligence. Again, this is a voluntary undertaking. This means that coverage is patchy and could lead to inefficiencies compared to the risk pooling model.

The modifications to insurance terms introduced under statute to deal with self-driving vehicles have the advantage that drivers must already carry compulsory insurance for risks to third parties.⁹⁹ In order for claimants to receive the protection of the risk pooling scheme, it would be necessary for one or more of the stakeholders to be under a statutory obligation to hold compulsory insurance for harms related to AIS use. Insurers for clinical malpractice are best placed to process claims directly with patients. Rules directing that clinical insurers are liable to pay claims in the first instance would mean that the patient would not need to identify multiple potential defendants in order to pursue a claim.¹⁰⁰

The AEV Act 2018 again provides a useful pattern to follow. In Section IV it was acknowledged that software normally requires revisions over time as and when issues are discovered. The legislation for self-driving cars provides for this by instituting a rule that

⁹⁷ Medical Act 1983, s 44C; Nursing and Midwifery Order 2001 (SI 2002/253), art 12A; General Medical Council, *Good medical practice* (2013) para 63 < <https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/good-medical-practice>> accessed 8 March 2020; Nursing and Midwifery Council, *The Code* (2015) para 12 <www.nmc.org.uk/standards/code/> accessed 8 March 2020

⁹⁸ NHS, *Claims Management* (19 July 2019) <<https://resolution.nhs.uk/services/claims-management/>> accessed 8 March 2020

⁹⁹ Automated and Electric Vehicles Act 2018, s 20, sch 1, paras 17-19

¹⁰⁰ *cf* Automated and Electric Vehicles Act 2018, s 2(1)

insurers may exclude liability where the software has not been updated at the time of the accident.¹⁰¹ If a similar rule were to be introduced into clinical malpractice schemes, this would have the effect of encouraging hospitals to keep AISs up to date, with the by-product that patient care would be continually at the forefront of scientific and technical capability. The SDC has a duty to ensure that their system is safe. If they do not, they may be obliged to recall it or make it safe.¹⁰² If the SDC fails to make a system safe, for example, via an update, they have breached a statutory duty. Once this update is available, the clinician cannot claim against the insurance scheme if they have not updated the system.

2. Insurance schemes for clinical malpractice should have powers to recover costs

Although the insurer is intended to be the sole point of contact for a claimant, risk pooling does not require that the insurer fund the full amount of compensation from the insurer's own reserves. The notion of risk pooling recognises that the totality of the damage is the result of numerous, sometimes undetectable, errors and mistakes. The actions or omissions of one party may indeed be the nearest cause to the damage in time and space, yet it is still legitimate to seek a contribution from co-defendants whose negligence also led to the occurrence of the damage.¹⁰³

This entitlement exists in statute;¹⁰⁴ a modified power also exists for the insurer of a self-driving vehicle to recover the costs of a claim award which has already been paid to the claimant.¹⁰⁵ In this sense, the insurer acts as a centralised claims administrator but can also investigate specific incidents in more depth. An individual claimant may only be able to amass enough evidence to merit a claim against some defendants but not others. Meanwhile, insurers

¹⁰¹ Automated and Electric Vehicles Act 2018, s 4(1)

¹⁰² Medical Devices Regulations 2002 (SI 2002/618), r 63 (5) (a) and (b)

¹⁰³ *Hughes v Williams* [2013] PIQR P17

¹⁰⁴ Civil Liability (Contribution) Act 1978

¹⁰⁵ Automated and Electric Vehicles Act 2018, s 5

can mobilise their institutional resources in preparing a claim. Moreover, they are incentivised to seek contributions from other parties as it can eliminate or reduce the burden on the insured community, thereby keeping premiums at competitive rates.

3. Several liability should be abolished

The power to recover costs or seek financial contribution to an award of damages rests on the model of joint liability, as was expressed in *Fairchild v Glenhaven Funeral Services Ltd*.¹⁰⁶ Joint liability is a legal fiction with a pragmatic purpose; it acknowledges that early and full compensation is crucial for the claimant to have the stability to start their journey of recovery as soon as practicable. Risk pooling is grounded in the idea that the stakeholders who intend to benefit from their participation in a risky activity ought to accept moral and legal responsibilities to the individuals and communities that may suffer the consequences of when things go awry.

The court in *Barker v Corus UK Ltd*¹⁰⁷ erred in making liability for ‘material increase in the risk’ ‘several’, or proportionate to the defendant’s share of the risk exposure. It is argued here that any party that negligently exposes a claimant to a risk of harm via an AI system used in health care should bear full responsibility for the damage if that risk materialises. If Parliament can recognise that the principle in *Barker* does injustice by mesothelioma claimants,¹⁰⁸ then surely it must appreciate that the risks of AI-driven healthcare, and indeed other risky activities, are so inextricably linked between parties that joint liability should be restored.

4. Patient consent should be restricted

For some activities, the nature of the risk involved is so great that the law has seen fit to restrain unwitting claimants from granting effective consent as this would bar them from compensation

¹⁰⁶ [2003] 1 AC 32

¹⁰⁷ [2006] 2 AC 572

¹⁰⁸ Compensation Act 2006, s 3

to which they might otherwise be entitled. For instance, it has been held that passengers are not able to form the necessary acceptance of risk required for the defence of *volenti non fit injuria* where the driver of the vehicle is intoxicated.¹⁰⁹ Statute now excludes this defence in its entirety in road traffic accidents, because all drivers are required to have insurance in order for the speedy administration of justice to claimants.¹¹⁰

A similar contention can be made in favour of patients who have been harmed as a result of defective AISs. As laypersons, patients cannot be expected to take the steps to safeguard their own interests, particularly when explanations of how some AISs operate are not readily forthcoming for commercial and technical reasons. Restricting the patient's ability to consent to treatment by defective AISs would provide enhanced protection for vulnerable claimants.

B. Advantages of risk pooling

Risk pooling would only be ethically and professionally sound if schemes are fit for use. Justice would not be achieved if the compulsory insurance component of risk pooling was used to rationalise the rapid deployment of AISs just because one could rely on insurance to cover the risks. Nevertheless, if the risk pooling approach were to be adopted, insurance companies would drive up safety standards in order to protect their business. For example, insurers would be reluctant to enter into commercial partnerships where proof of compliance with technical standards as well as NICE and NHSX (2019) guidance is absent.

Principle 10 of the NHSX Code of Conduct for Data-driven Health and Care Technology demands that 'clear terms around who bears the liability should be established',¹¹¹ and it seems

¹⁰⁹ *Dann v Hamilton* [1939] 1 KB 509

¹¹⁰ Road Traffic Act 1988, s 149

¹¹¹ NHSX, *Code of conduct for data-driven health and care technology* (18 July 2019)

<www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology> 8 March 2020

that NHSX envisage that this will be the outcome of contractual negotiations between healthcare providers and SDCs.

As standards become more detailed and based on empirical evidence, premiums could conceivably reduce (though this is essentially a question for actuaries, not lawyers). If the cost of insurance is prohibitive because the risks are very high, this financial inconvenience should be welcomed as it will prevent harmful systems from being used in clinical practice at a time when those risks cannot be effectively mitigated.

VI. CONSIDERATIONS FOR THE FUTURE AND CONCLUSIONS

The law addressing liability in the use of artificial intelligence for clinical decision-making is complex and ill-defined for non-legally minded stakeholders. *Novus actus interveniens* seems to protect the SDC whilst leaving the clinician vulnerable to negligence claims. It is difficult to assess with any certainty which claims would succeed or fail in this area as caselaw on the points raised in this article is non-existent and existing legal frameworks might not be desirable.¹¹² There is an opportunity to take a proactive approach to ensure patient safety and compensation, thus, it is logical to have the insurance, standards, and regulation in place.

If SDCs do not wish to risk caselaw determining whether or not they have a duty of care to the patients which their systems are trying to help, there is merit in their proactively embracing their potential duty of care to patients and to adopt industry agreed standards and codes of practice.

No ethical approval was required for this article. The authors declare no conflicts of interest. The authors received no funding for the production of this work.

¹¹² Select Committee on Artificial Intelligence, *AI in the UK: ready, willing and able?* (HL 2017-19, 100) 96